

# Capitolo 7

## Autovalori e autovettori

### 7.1 Definizione, esempi e applicazioni

**Definizione 7.1.** Sia  $A \in M_n(\mathbb{K})$  una matrice quadrata di ordine  $n$  a entrate in un campo  $\mathbb{K}$ . Un elemento  $\lambda \in \mathbb{K}$  si dice **autovalore di  $A$**  se esiste una  $n$ -upla non nulla  $0 \neq (x_1, x_2, \dots, x_n) \in \mathbb{K}^n$  tale che

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \quad (7.1)$$

In tal caso il vettore  $(x_1, x_2, \dots, x_n)$  è detto **autovettore di  $A$  relativo all'autovalore  $\lambda$** .

Se  $x$  denota la  $n$ -upla scritta in colonna, scriveremo la (7.1) in modo più conciso come prodotto di matrici

$$Ax = \lambda x.$$

Un autovettore è quindi una  $n$ -upla che viene mandata in un suo multiplo quando le viene applicata la matrice  $A$ . Ad esempio, se  $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$  e  $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ , si ha

$$Ax = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 4x$$

e quindi  $x$  è un autovettore di  $A$  associato all'autovalore  $\lambda = 4$ .

Per contro il vettore  $x = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$  non è un autovettore di  $A$  in quanto

$$Ax = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \end{pmatrix}$$

e (5, 7) non è multiplo di (2, 1).

Prima di vedere come si calcolano autovalori e autovettori di una matrice data, citiamo alcune applicazioni in fisica, matematica ed informatica.

- (1) Una massa collegata a una molla, se spostata dalla posizione di equilibrio, inizia a oscillare con una certa frequenza, dipendente dalla massa stessa e dalle caratteristiche fisiche della molla. Se invece abbiamo più masse collegate tra loro tramite delle molle, l'oscillazione di ognuna di esse dipende anche dalla posizione e dal movimento delle altre masse del sistema e bisogna quindi tener conto di tutte le interazioni reciproche. Si mostra che tali interazioni possono essere descritte mediante una matrice quadrata, e gli autovalori di questa matrice ci danno le cosiddette *frequenze naturali di oscillazione del sistema*. Il modo in cui il sistema oscilla è dato da una combinazione di queste frequenze tramite una formula che coinvolge anche gli autovettori corrispondenti.
- (2) Abbiamo visto nel Capitolo 4 come, fissato un sistema di riferimento nello spazio tridimensionale, una generica equazione  $Ax + By + Cz = D$  di primo grado in tre incognite rappresenti un piano, nel senso che i punti appartenenti al piano sono tutti e soli quelli le cui coordinate  $(x, y, z)$  verificano l'equazione.

Ci chiediamo ora cosa sia rappresentato da una generica equazione di secondo grado nelle tre incognite date, ovvero

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + L = 0. \quad (7.2)$$

Si può dimostrare che le equazioni di questo tipo rappresentano superfici dette *quadriche* (tra le quali rientrano sfere, ellissoidi, iperboloidi, coni, cilindri) e che, per capire quale tra queste superfici rappresenti l'equazione (7.2), bisogna calcolare gli autovalori di una matrice di ordine 4 costruita opportunamente dai coefficienti  $A, B, C, \dots, L$ :

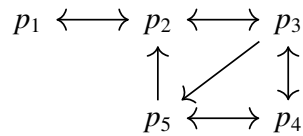
$$\begin{pmatrix} A & D & E & G \\ D & B & F & H \\ E & F & C & I \\ G & H & I & L \end{pmatrix}$$

Inoltre, gli autovettori forniscono informazioni sulla posizione di tale superficie nello spazio, dicendoci ad esempio quali sono i suoi assi di simmetria.

- (3) Come ultima, importante applicazione della teoria degli autovalori e autovettori, vediamo che tali nozioni sono alla base dell'algoritmo con cui Google ordina le pagine internet per importanza. Visto l'interesse di tale esempio per l'informatica, diamo qualche dettaglio in più rispetto agli esempi precedenti.

La rete internet consiste di un numero (enorme ma finito) di pagine  $p_1, p_2, \dots, p_N$  ciascuna delle quali può avere uno o più link verso le altre pagine. Ad esempio,

supponiamo che internet sia formato solo da 5 pagine  $p_1, p_2, p_3, p_4, p_5$ . Lo schema seguente



rappresenta una rete nella quale la notazione  $p_i \rightarrow p_j$  significa che c'è un link dalla pagina  $p_i$  alla pagina  $p_j$  e una doppia freccia  $p_i \leftrightarrow p_j$  significa che esiste sia un link da  $p_i$  verso  $p_j$  che viceversa.

L'idea, abbastanza naturale, dell'algoritmo di Google, è che una pagina  $p$  è tanto più importante quanto maggiore è il numero di pagine che hanno un link verso  $p$ . Tuttavia i link non hanno tutti lo stesso valore: questo dipenderà a sua volta dall'importanza della pagina da cui proviene.

Per formalizzare quest'idea e dare delle formule rigorose, iniziamo con il denotare con  $|p|$  il numero di link che vanno da una pagina  $p$  verso altre pagine. Ad esempio, facendo riferimento alla rete di 5 pagine rappresentata nel disegno precedente, si ha

$$|p_1| = 1, |p_2| = 2, |p_3| = 3, |p_4| = 2, |p_5| = 2.$$

Ora, per ogni pagina  $p$  definiamo l'importanza o il *rank* di  $p$  come il numero positivo  $r(p)$  che soddisfa l'uguaglianza

$$r(p) = \sum_{q \rightarrow p} \frac{r(q)}{|q|} \quad (7.3)$$

dove la somma è presa su tutte le pagine  $q$  che hanno un link verso  $p$ . Tale formula formalizza matematicamente l'idea di importanza di una pagina che abbiamo descritto intuitivamente sopra. Infatti, maggiore è il numero di link che riceve  $p$  maggiore sarà il risultato della sommatoria (avendo più addendi). Inoltre ognuno degli addendi,  $\frac{r(q)}{|q|}$ , è dato dall'importanza  $r(q)$  della pagina  $q$  (quanto più  $q$  è importante, tanto più sarà importante  $p$ , verso la quale  $q$  ha un link), ma diviso per il numero  $|q|$  totale dei link che  $q$  ha verso le altre pagine della rete (se  $q$  ha link verso più pagine, ciascuna di queste pagine "eredita" solo una frazione dell'importanza di  $q$  nel calcolo della propria importanza).

Il punto è che la formula (7.3) definisce l'importanza  $r(p)$  di  $p$  in funzione delle importanze  $r(q)$  di altre pagine per calcolare la quale dovremmo conoscere anche  $r(p)$ . Questo apparente circolo vizioso si risolve riformulando la definizione (7.3) proprio in termini di autovalori e autovettori.

Più precisamente, costruiamo una matrice  $A$  quadrata di ordine  $N$ , dove  $N$  è il numero delle pagine della rete, e che rappresenti la rete nel modo seguente. Definiamo le entrate  $a_{ij}$  di  $A$  come

$$a_{ij} = \begin{cases} 0 & \text{se non esiste nessun link da } p_j \text{ a } p_i, \\ \frac{1}{|p_j|} & \text{se esiste un link da } p_j \text{ a } p_i. \end{cases}$$

Ad esempio, si vede che per la rete di 5 pagine rappresentata nel disegno sopra si ha

$$A = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 1 & 0 & \frac{1}{3} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{3} & \frac{1}{2} & 0 \end{pmatrix} \quad (7.4)$$

In pratica, nella  $j$ -esima colonna di tale matrice leggiamo verso quali altre pagine ci sono link. Ad esempio, la prima colonna di  $A$  ha l'unica entrata non nulla in seconda posizione, il che corrisponde al fatto che  $p_1$  ha un link solo verso  $p_2$ ; la seconda colonna ha due entrate non nulle, la prima e la terza, che corrisponde al fatto che da  $p_2$  partono due link, uno verso  $p_1$  e un altro verso  $p_3$ , e il valore assegnato è  $\frac{1}{2}$ , e così via.

Ora mostriamo che trovare l'importanza delle pagine equivale a risolvere un problema di autovettori per  $A$ . Più precisamente, mostriamo che la (7.3) corrisponde all'uguaglianza

$$Ar = r \quad (7.5)$$

dove

$$r = \begin{pmatrix} r(p_1) \\ r(p_2) \\ \vdots \\ r(p_N) \end{pmatrix}$$

è la  $n$ -upla le cui componenti sono i rank delle pagine, che dobbiamo determinare. In altre parole, in base alla (7.5), tale  $n$ -upla è un autovettore di  $A$  relativo all'autovalore  $\lambda = 1$ .

Infatti, la (7.5) ci dice che la  $j$ -esima componente  $(Ar)_j$  di  $Ar$  deve essere uguale alla  $j$ -esima componente di  $r$ , cioè  $r(p_j)$ . Ma la  $j$ -esima componente  $(Ar)_j$  del prodotto di  $A$  per  $r$  si ottiene dal prodotto righe per colonne della  $j$ -esima riga di  $A$  per  $r$ , cioè

$$(Ar)_j = a_{j1}r(p_1) + a_{j2}r(p_2) + \cdots + a_{jN}r(p_N) = \sum_{k=1}^N a_{jk}r(p_k). \quad (7.6)$$

Tuttavia, in base alla definizione della matrice  $A$ , in tale sommatoria un'entrata  $a_{jk}$  è diversa da zero solo se esiste un link da  $p_k$  verso  $p_j$ . Quindi, tenendo conto solo degli addendi non nulli, possiamo riscrivere la (7.6) come

$$(Ar)_j = \sum_{p_k \rightarrow p_j} a_{jk} r(p_k).$$

Ma allora, ricordando che, se esiste un link da  $p_k$  verso  $p_j$  allora  $a_{jk} = \frac{1}{|p_k|}$ , otteniamo

$$(Ar)_j = \sum_{p_k \rightarrow p_j} \frac{1}{|p_k|} r(p_k).$$

Quindi per la formula (7.3), si vede che il secondo membro di quest'ultima uguaglianza è esattamente  $r(p_j)$ , il che mostra che  $(Ar)_j = r(p_j)$ , come volevamo.

Ad esempio, per la nostra rete di 5 pagine, rappresentata dalla matrice  $A$  data in (7.4), si può dimostrare che

$$r = (r(p_1), r(p_2), r(p_3), r(p_4), r(p_5)) = \left( \frac{2}{13}, \frac{4}{13}, \frac{3}{13}, \frac{2}{13}, \frac{2}{13} \right)$$

il che ci dice in particolare che  $p_2$  è la pagina più importante (il che è ragionevole in quanto è la pagina che riceve più link), seguita da  $p_3$ , che riceve 2 link esattamente come  $p_4$  e  $p_5$ , ma uno dei suoi link viene da  $p_2$  che è la pagina più importante.

Come vedremo nel prossimo paragrafo, non è detto che una matrice ammetta l'autovalore  $\lambda = 1$ , quindi potrebbe non esistere un vettore  $r$  che soddisfi  $Ar = r$  le cui componenti siano date dalle importanze delle pagine così come le abbiamo definite tramite la (7.3). In tal caso, la matrice  $A$  che descrive la rete va leggermente modificata in modo che tale ipotesi sia soddisfatta. Un ulteriore problema è quello relativo all'ordine della matrice, uguale al numero delle pagine della rete, che nella realtà è un numero enorme, il che rende praticamente impossibile il calcolo esatto di autovalori e autovettori. Si ricorre quindi a soluzioni approssimate.<sup>1</sup>

*Osservazione 7.2.* Gli esempi precedenti ci danno un'ulteriore conferma delle numerose applicazioni che le matrici hanno nei più svariati contesti: non solo esse possono rappresentare sistemi di equazioni o trasformazioni geometriche, come abbiamo visto nei capitoli precedenti, ma, come si vede sia dall'esempio dei sistemi vibranti che da quello di Google, sono lo strumento matematico ideale per rappresentare sistemi costituiti da un numero finito di elementi che interagiscono tra loro in qualche modo.

<sup>1</sup>Per ulteriori dettagli rimandiamo alla lettura dell'articolo originale di Larry Page e Sergei Brin, fondatori di Google: *The Anatomy of a Large-Scale Hypertextual Web Search Engine* (1998) Seventh International World-Wide Web Conference (WWW 1998), April 14-18, 1998, Brisbane, Australia.

## 7.2 Calcolo di autovalori e autovettori

Come abbiamo visto nella Definizione 7.1 trovare gli autovettori di una matrice  $A$  significa trovare gli  $x$  diversi dalla  $n$ -upla nulla tali che  $Ax = \lambda x$  per qualche  $\lambda \in \mathbb{K}$ . Vediamo ora come questa uguaglianza può essere riscritta come un sistema di equazioni lineari al quale applicare la teoria vista nei capitoli precedenti.

Come abbiamo visto alla fine del capitolo precedente, sfruttando le proprietà delle matrici, l'uguaglianza  $Ax = \lambda x$  può essere riscritta come

$$Ax - \lambda x = 0 \quad (7.7)$$

dove stiamo indicando con  $0$  la  $n$ -upla nulla, scritta in colonna. Ora, per poter mettere in evidenza  $x$  nella (7.7), sfruttando la proprietà distributiva  $AB + AC = A(B + C)$  del prodotto di matrici, usiamo prima  $x = \text{Id}_n x$  ottenendo

$$Ax - \lambda \text{Id}_n x = 0$$

da cui possiamo ora mettere in evidenza

$$(A - \lambda \text{Id}_n)x = 0. \quad (7.8)$$

Questa uguaglianza rappresenta un sistema omogeneo di  $n$  equazioni lineari in  $n$  incognite (dove  $n$  è l'ordine della matrice  $A$ ), e trovare gli autovettori  $x$  di  $A$  equivale a trovare le soluzioni  $x \neq 0$  di tale sistema.

Ricordiamo che un sistema omogeneo è sempre compatibile e ammette almeno la soluzione nulla. Il sistema (7.8) con  $n$  incognite ammette solo la soluzione nulla se e solo se il rango della sua matrice dei coefficienti è uguale a  $n$ .

Dal momento che siamo interessati all'esistenza di altre soluzioni oltre a quella nulla, la condizione necessaria e sufficiente per l'esistenza di tali soluzioni è quindi che il rango della matrice dei coefficienti  $A - \lambda \text{Id}_n$  del sistema (7.8) sia minore di  $n$ .

Tale sistema dipende però da  $\lambda$  e esso ammetterà soluzioni non nulle solo per certi valori di  $\lambda$ . Quello che faremo sarà quindi prima determinare i  $\lambda \in \mathbb{K}$  per cui tali soluzioni esistono, e poi risolvere il sistema solo in corrispondenza di quei valori. Ma, come abbiamo visto nel capitolo precedente, una matrice quadrata di ordine  $n$  ha rango minore di  $n$  se e solo se il suo determinante si annulla. In conclusione, gli autovalori di  $A$  sono i  $\lambda \in \mathbb{K}$  che soddisfano

$$\det(A - \lambda \text{Id}_n) = 0 \quad (7.9)$$

e gli autovettori di  $A$  si trovano come soluzioni del sistema (7.8) per i rispettivi autovalori.

**Definizione 7.3.** Come vedremo negli esempi, il determinante della matrice  $A - \lambda \text{Id}_n$  è un polinomio di grado  $n$  (uguale all'ordine della matrice) in  $\lambda$  che si chiama **polinomio caratteristico di  $A$** , e l'equazione (7.9) si dice **equazione caratteristica**.

*Osservazione 7.4.* Si noti che

$$\begin{aligned} A - \lambda \text{Id}_n &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \\ &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} - \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix} = \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix} \end{aligned}$$

ovvero  $A - \lambda \text{Id}_n$  è semplicemente la matrice  $A$  in cui abbiamo sottratto  $\lambda$  alle entrate della diagonale.

*Esempio 7.5.* Determiniamo autovalori e autovettori della seguente matrice a entrate reali

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}.$$

Iniziamo col calcolare il polinomio caratteristico di  $A$ . Si ha

$$\det(A - \lambda \text{Id}_2) = \det \begin{pmatrix} 1 - \lambda & 3 \\ 3 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2 - 9 = \lambda^2 - 2\lambda - 8.$$

L'equazione caratteristica, le cui soluzioni reali sono gli autovalori di  $A$ , è quindi

$$\lambda^2 - 2\lambda - 8 = 0.$$

Applicando la formula risolutiva

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

per la generica equazione di secondo grado  $ax^2 + bx + c = 0$ , si trova subito  $\lambda = 4$  e  $\lambda = -2$ , che sono quindi gli autovalori di  $A$ .

Per trovare gli autovettori relativi a  $\lambda = 4$ , basta risolvere il sistema omogeneo che ha  $A - 4\text{Id}_2$  come matrice dei coefficienti. Analogamente, per trovare gli autovettori relativi a  $\lambda = -2$ , basta risolvere il sistema omogeneo che ha  $A + 2\text{Id}_2$  come matrice dei coefficienti.

Per  $\lambda = 4$  si ha  $A - \lambda \text{Id}_2 = \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix}$ . Come si vede, le righe rappresentano le equazioni equivalenti (una opposta dell'altra)  $-3x_1 + 3x_2 = 0$  e  $3x_1 - 3x_2 = 0$ . Dividendo per 3, si ha quindi l'unica equazione  $-x_1 + x_2 = 0$ . Posto  $x_2 = t$ , si trova  $x_1 = x_2 = t$ . Quindi le soluzioni del sistema  $(A - 4\text{Id}_2)x = 0$ , ovvero gli autovettori di  $A$  relativi all'autovalore 4, sono esattamente tutti i vettori del tipo  $(t, t)$ , al variare di  $t \in \mathbb{R}$ . Ad esempio, per  $t = 1$  si ritrova proprio l'autovettore visto subito dopo la Definizione [7.1](#)

Analogamente, per  $\lambda = -2$  si ha  $A - \lambda \text{Id}_2 = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$ , quindi le righe rappresentano la stessa equazione  $3x_1 + 3x_2 = 0$ . Dividendo per 3 e posto  $x_2 = t$ , si trova  $x_1 = -x_2 = -t$ . Quindi le soluzioni del sistema  $(A + 2\text{Id}_2)x = 0$ , ovvero gli autovettori di  $A$  relativi all'autovalore  $-2$ , sono esattamente tutti i vettori del tipo  $(-t, t)$ , al variare di  $t \in \mathbb{R}$ .

*Esempio 7.6.* Sia

$$A = \begin{pmatrix} 5 & 1 & -1 \\ 4 & 5 & -2 \\ 2 & 1 & 2 \end{pmatrix}.$$

Per calcolare il polinomio caratteristico svolgiamo

$$\det(A - \lambda \text{Id}_3) = \det \begin{pmatrix} 5 - \lambda & 1 & -1 \\ 4 & 5 - \lambda & -2 \\ 2 & 1 & 2 - \lambda \end{pmatrix} =$$

(sviluppando secondo Laplace rispetto alla prima riga)

$$\begin{aligned} &= (5 - \lambda) \det \begin{pmatrix} 5 - \lambda & -2 \\ 1 & 2 - \lambda \end{pmatrix} - \det \begin{pmatrix} 4 & -2 \\ 2 & 2 - \lambda \end{pmatrix} - \det \begin{pmatrix} 4 & 5 - \lambda \\ 2 & 1 \end{pmatrix} = \\ &(5 - \lambda)[(5 - \lambda)(2 - \lambda) + 2] - [4(2 - \lambda) + 4] - [4 - 2(5 - \lambda)] \end{aligned}$$

ovvero, svolgendo i calcoli,

$$-\lambda^3 + 12\lambda^2 - 45\lambda + 54.$$

Si noti che, come abbiamo anticipato sopra, il polinomio caratteristico è un polinomio di grado uguale all'ordine della matrice, ovvero 3.

Per trovare le *radici* del polinomio caratteristico, ovvero le soluzioni dell'equazione caratteristica

$$-\lambda^3 + 12\lambda^2 - 45\lambda + 54 = 0 \tag{7.10}$$

possiamo usare il seguente risultato algebrico.

*Lemma 7.7.* *Se un polinomio a coefficienti interi ammette una radice razionale, questa è necessariamente della forma  $\frac{a}{b}$  dove  $a$  è un divisore del termine noto e  $b$  è un divisore del coefficiente del termine di grado massimo.*

Nel caso della nostra equazione (7.10), essendo  $-1$  il coefficiente del termine di grado massimo (e non avendo quindi altri divisori oltre a  $\pm 1$ ) le eventuali soluzioni razionali, se esistono, sono tra i divisori del termine noto 54, ovvero  $\pm 1, \pm 2, \pm 3, \pm 6, \pm 9, \pm 18, \pm 27, \pm 54$ .

Ad esempio, sostituendo  $\lambda = 1$  nella (7.10) si trova

$$-1^3 + 12 \cdot 1^2 - 45 \cdot 1 + 54 = 20 \neq 0$$

e quindi 1 non è soluzione dell'equazione. D'altra parte sostituendo  $\lambda = 3$  si ottiene

$$-3^3 + 12 \cdot 3^2 - 45 \cdot 3 + 54 = -27 + 108 - 135 + 54 = 0.$$

Quindi  $\lambda = 3$  è soluzione dell'equazione caratteristica (ed è quindi un autovalore di  $A$ ). Una volta trovata una soluzione di un'equazione polinomiale, le altre possono essere trovate grazie al seguente risultato, che ci permette di ridurre l'equazione a una di grado più basso.

*Teorema 7.8 (Teorema di Ruffini). Se un polinomio  $P$  di grado  $n$  in  $\lambda$  ammette una radice  $\lambda_0$ , allora esso è divisibile per  $\lambda - \lambda_0$ . Ovvero  $P$  si decompone come prodotto  $P = (\lambda - \lambda_0)P'$  dove  $P'$  è un polinomio di grado  $n - 1$  in  $\lambda$ .*

A questo punto, l'equazione  $P = 0$  equivale a  $(\lambda - \lambda_0)P' = 0$ . Questa, per la legge di annullamento del prodotto, può essere verificata solo in due casi:  $\lambda - \lambda_0 = 0$ , cioè  $\lambda = \lambda_0$  (che ci dà la soluzione che già conoscevamo) oppure  $P' = 0$ . Ci siamo quindi ridotti a risolvere un'equazione di grado più basso rispetto a quella iniziale.

Per applicare nella pratica questo metodo, bisogna conoscere un modo per trovare il fattore  $P'$  nella decomposizione  $P = (\lambda - \lambda_0)P'$ . A questo scopo, si usa un algoritmo che descriviamo tramite l'esempio della nostra equazione (7.10), per cui  $P = -\lambda^3 + 12\lambda^2 - 45\lambda + 54$  e  $\lambda_0 = 3$ . Si inizia riportando, come nello schema seguente, i coefficienti che moltiplicano i monomi che compongono il polinomio (da quello di grado più alto a quello di grado più basso) e in basso a sinistra la radice già trovata, ovvero  $\lambda = 3$ :

$$\begin{array}{r|rrr|r} & -1 & 12 & -45 & 54 \\ 3 & & & & \\ \hline & & & & \end{array}$$

Ora “abbassiamo”, riportandolo nella riga in basso, il primo coefficiente del polinomio

$$\begin{array}{r|rrr|r} & -1 & 12 & -45 & 54 \\ 3 & & & & \\ \hline & -1 & & & \end{array}$$

e moltiplichiamo il  $-1$  così abbassato per la radice  $3$  e riportiamo il risultato sotto il secondo coefficiente del polinomio

$$\begin{array}{c|ccc|c} & -1 & 12 & -45 & 54 \\ 3 & & -3 & & \\ \hline & -1 & & & \end{array}$$

Ora, sommiamo  $12$  e  $-3$  e riportiamo il risultato nella riga in basso

$$\begin{array}{c|ccc|c} & -1 & 12 & -45 & 54 \\ 3 & & -3 & & \\ \hline & -1 & 9 & & \end{array}$$

A questo punto iteriamo il procedimento. Così come prima abbiamo moltiplicato  $-1$  per la radice  $\lambda = 3$ , ora moltiplichiamo il  $9$  appena aggiunto nella riga in basso per  $\lambda = 3$ , scriviamo il risultato nella colonna successiva sotto il  $-45$ , sommiamo e riportiamo nella riga in basso:

$$\begin{array}{c|ccc|c} & -1 & 12 & -45 & 54 \\ 3 & & -3 & 27 & \\ \hline & -1 & 9 & -18 & \end{array}$$

Come prevede l'algoritmo iteriamo il procedimento: moltiplichiamo il  $-18$  appena aggiunto nella riga in basso per  $\lambda = 3$ , scriviamo il risultato nella colonna successiva sotto il  $54$ , sommiamo e riportiamo nella riga in basso.

$$\begin{array}{c|ccc|c} & -1 & 12 & -45 & 54 \\ 3 & & -3 & 27 & -54 \\ \hline & -1 & 9 & -18 & 0 \end{array}$$

L'ultimo  $0$  riportato è la conferma del fatto che abbiamo svolto i calcoli correttamente, e i tre coefficienti  $-1, 9, -18$  ottenuti prima dello zero sono proprio i coefficienti del polinomio  $P'$  (di grado  $2$ ) tale che  $P = (\lambda - 3)P'$ , ordinati dal termine di grado più alto fino al termine noto. Possiamo quindi scrivere la decomposizione in fattori di  $P$  come

$$-\lambda^3 + 12\lambda^2 - 45\lambda + 54 = (\lambda - 3)(-\lambda^2 + 9\lambda - 18).$$

A questo punto, come abbiamo detto sopra, l'equazione caratteristica ha come soluzioni, oltre a  $\lambda = 3$ , anche i  $\lambda$  che annullano l'altro fattore, ovvero tali che  $-\lambda^2 + 9\lambda - 18 = 0$ . Risolvendo con la formula risolutiva per le equazioni di secondo grado si trova che

questa equazione ha come soluzioni  $\lambda = 3$  e  $\lambda = 6$ .

Quindi gli autovalori della matrice  $A$  sono  $\lambda = 3$  e  $\lambda = 6$ , con  $\lambda = 3$  che si ripete due volte come soluzione dell'equazione.

Possiamo ora determinare gli autovettori corrispondenti. Come abbiamo visto nella (7.8), dobbiamo risolvere il sistema omogeneo che ha  $A - \lambda \text{Id}_3$  come matrice dei coefficienti, prima con  $\lambda = 3$  e poi con  $\lambda = 6$ . Per  $\lambda = 3$  si ha

$$A - \lambda \text{Id}_3 = \begin{pmatrix} 2 & 1 & -1 \\ 4 & 2 & -2 \\ 2 & 1 & -1 \end{pmatrix}.$$

Riducendo tale matrice a gradini<sup>2</sup> si trova subito

$$\begin{pmatrix} 2 & 1 & -1 \\ 4 & 2 & -2 \\ 2 & 1 & -1 \end{pmatrix} \xrightarrow[\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - R_1}]{} \begin{pmatrix} 2 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

e quindi il sistema  $(A - \lambda \text{Id}_3)x = 0$  si riduce all'unica equazione  $2x_1 + x_2 - x_3 = 0$ . Posti allora  $x_2 = t$  e  $x_3 = s$ , si trova  $2x_1 = -t + s$ , ovvero  $x_1 = \frac{-t+s}{2}$ . Le soluzioni del sistema, ovvero gli autovettori di  $A$  relativi all'autovalore  $\lambda = 3$ , sono quindi tutte le terne del tipo

$$\left( \frac{-t+s}{2}, t, s \right)$$

al variare di  $t, s \in \mathbb{R}$ .

Analogamente, per  $\lambda = 6$  si ha

$$A - \lambda \text{Id}_3 = \begin{pmatrix} -1 & 1 & -1 \\ 4 & -1 & -2 \\ 2 & 1 & -4 \end{pmatrix}.$$

Riducendo tale matrice a gradini si trova

$$\begin{pmatrix} -1 & 1 & -1 \\ 4 & -1 & -2 \\ 2 & 1 & -4 \end{pmatrix} \xrightarrow[\substack{R_2 \rightarrow R_2 + 4R_1 \\ R_3 \rightarrow R_3 + 2R_1}]{} \begin{pmatrix} -1 & 1 & -1 \\ 0 & 3 & -6 \\ 0 & 3 & -6 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - R_2} \begin{pmatrix} -1 & 1 & -1 \\ 0 & 3 & -6 \\ 0 & 0 & 0 \end{pmatrix}$$

Dividendo inoltre la seconda riga per 3, vediamo che il sistema  $(A - \lambda \text{Id}_3)x = 0$  si riduce a

$$\begin{cases} -x_1 + x_2 - x_3 = 0 \\ x_2 - 2x_3 = 0 \end{cases}$$

<sup>2</sup>Ma basta anche osservare che la seconda e la terza riga sono dipendenti dalla prima.

Posto  $x_3 = t$ , dalla seconda equazione si trova  $x_2 = 2t$ , e sostituendo nella prima  $-x_1 + 2t - t = 0$ , ovvero  $x_1 = t$ . Le soluzioni del sistema, ovvero gli autovettori di  $A$  relativi all'autovalore  $\lambda = 6$ , sono quindi tutte le terne del tipo

$$(t, 2t, t)$$

al variare di  $t \in \mathbb{R}$ .

Questo esempio mostra che un autovalore  $\lambda_0$  può presentarsi più volte come soluzione dell'equazione caratteristica e che gli autovettori relativi ad esso possono dipendere da più parametri.

**Definizione 7.9.** Sia  $A$  una matrice quadrata di ordine  $n$  e  $\lambda_0$  un suo autovalore. La **molteplicità algebrica** di  $\lambda_0$  è il numero di volte che  $(\lambda - \lambda_0)$  divide il polinomio caratteristico  $P = \det(A - \lambda \text{Id}_n)$ . Si chiama invece **molteplicità geometrica** di  $\lambda_0$  il numero di parametri indipendenti che descrivono l'insieme dei suoi autovettori.

Nell'Esempio 7.6  $\lambda = 3$  ha molteplicità sia algebrica che geometrica pari a 2, mentre l'autovalore  $\lambda = 6$  ha molteplicità algebrica e geometrica entrambe uguali a 1.

In generale, non è detto che molteplicità algebrica e geometrica di un autovalore siano uguali. Si può anzi dimostrare il seguente risultato.

**Teorema 7.10.** *La molteplicità geometrica di un autovalore è sempre minore o uguale alla sua molteplicità algebrica.*

*Osservazione 7.11.* La molteplicità geometrica di un autovalore  $\lambda_0$  è in effetti uguale a  $n - r$ , dove  $n$  è l'ordine della matrice e  $r$  è il rango della matrice  $A - \lambda_0 \text{Id}_n$ . Infatti, come sappiamo dalla Proposizione 5.17 il sistema  $(A - \lambda_0 \text{Id}_n)x = 0$  ha proprio  $\infty^{n-r}$  soluzioni, ovvero le sue soluzioni (che ci danno gli autovettori di  $A$  relativi all'autovalore  $\lambda_0$ ) si scrivono in funzione di  $n - r$  parametri. La molteplicità geometrica di un autovalore è quindi sempre maggiore o uguale a 1. In particolare se un autovalore ha molteplicità algebrica 1, allora, per il Teorema 7.10, anche la sua molteplicità geometrica sarà 1.

*Esempio 7.12.* Sia

$$A = \begin{pmatrix} 3 & \frac{1}{2} & 2 \\ -1 & 1 & -1 \\ -1 & -\frac{1}{2} & 0 \end{pmatrix}.$$

Il polinomio caratteristico è dato da

$$\det(A - \lambda \text{Id}_3) = \det \begin{pmatrix} 3 - \lambda & \frac{1}{2} & 2 \\ -1 & 1 - \lambda & -1 \\ -1 & -\frac{1}{2} & -\lambda \end{pmatrix} =$$

sviluppando secondo Laplace rispetto alla prima riga

$$= (3 - \lambda) \det \begin{pmatrix} 1 - \lambda & -1 \\ -\frac{1}{2} & -\lambda \end{pmatrix} - \frac{1}{2} \det \begin{pmatrix} -1 & -1 \\ -1 & -\lambda \end{pmatrix} + 2 \det \begin{pmatrix} -1 & 1 - \lambda \\ -1 & -\frac{1}{2} \end{pmatrix} =$$

e svolgendo i calcoli si ottiene

$$-\lambda^3 + 4\lambda^2 - 5\lambda + 2$$

Per il Lemma 7.7 le eventuali radici razionali vanno cercate tra i divisori del termine noto fratto i divisori del termine di grado massimo, quindi  $\pm 1, \pm 2$ .

Sostituendo  $\lambda = 1$  si trova

$$-1 + 4 + 2 - 5 = 0$$

ovvero  $\lambda = 1$  è soluzione dell'equazione caratteristica (ed è quindi un autovalore di  $A$ ). Per determinare le altre soluzioni, applichiamo l'algoritmo descritto nell'Esempio 7.6. Disponiamo nello schema già visto i coefficienti che moltiplicano i monomi che compongono il polinomio, da quello di grado più alto a quello di grado più basso, e in basso a sinistra la radice già trovata, ovvero  $\lambda = 1$ :

$$\begin{array}{r|rrr|r} & -1 & 4 & -5 & 2 \\ 1 & & & & \\ \hline & & & & \end{array}$$

Riportiamo nella riga in basso il primo coefficiente del polinomio

$$\begin{array}{r|rrr|r} & -1 & 4 & -5 & 2 \\ 1 & & & & \\ \hline & -1 & & & \end{array}$$

Moltiplichiamo il  $-1$  così abbassato per la radice 1 e riportiamo il risultato sotto il secondo coefficiente del polinomio

$$\begin{array}{r|rrr|r} & -1 & 4 & -5 & 2 \\ 1 & & -1 & & \\ \hline & -1 & & & \end{array}$$

Ora, sommiamo 4 e  $-1$  e riportiamo il risultato nella riga in basso

$$\begin{array}{ccc|c} & -1 & 4 & -5 & 2 \\ & & & & \\ 1 & & -1 & & \\ \hline & -1 & 3 & & \end{array}$$

Continuiamo ad applicare l'algoritmo come descritto nell'esempio precedente:

$$\begin{array}{ccc|c} & -1 & 4 & -5 & 2 \\ & & & & \\ 1 & & -1 & & \\ \hline & -1 & 3 & & \end{array} \longrightarrow \begin{array}{ccc|c} & -1 & 4 & -5 & 2 \\ & & & & \\ 1 & & -1 & 3 & \\ \hline & -1 & 3 & & \end{array} \longrightarrow$$

$$\longrightarrow \begin{array}{ccc|c} & -1 & 4 & -5 & 2 \\ & & & & \\ 1 & & -1 & 3 & \\ \hline & -1 & 3 & -2 & \end{array} \longrightarrow \begin{array}{ccc|c} & -1 & 4 & -5 & 2 \\ & & & & \\ 1 & & -1 & 3 & -2 \\ \hline & -1 & 3 & -2 & 0 \end{array}$$

Come abbiamo detto, l'ultimo 0 comparso è la conferma del fatto che abbiamo svolto i calcoli correttamente. I tre coefficienti  $-1, 3, -2$  ottenuti prima dello zero sono proprio i coefficienti del polinomio  $P'$  di grado 2 tale che  $P = (\lambda - 1)P'$ , ordinati dal termine di grado più alto fino al termine noto. Si ha quindi

$$-\lambda^3 + 4\lambda^2 - 5\lambda + 2 = (\lambda - 1)(-\lambda^2 + 3\lambda - 2).$$

A questo punto per trovare le altre radici del polinomio dobbiamo risolvere  $-\lambda^2 + 3\lambda - 2 = 0$ . Utilizzando la formula risolutiva per le equazioni di secondo grado si trovano le soluzioni  $\lambda = 1$  e  $\lambda = 2$ .

Concludiamo che gli autovalori della matrice  $A$  sono  $\lambda = 2$  e  $\lambda = 1$ , con quest'ultimo che si ripete quindi due volte come soluzione dell'equazione (in base alla Definizione [7.9](#) esso ha quindi molteplicità algebrica 2).

Possiamo ora determinare gli autovettori corrispondenti. Come abbiamo visto nella [\(7.8\)](#), dobbiamo risolvere il sistema omogeneo che ha  $A - \lambda \text{Id}_3$  come matrice dei coefficienti, prima con  $\lambda = 1$  e poi con  $\lambda = 2$ .

Per  $\lambda = 1$  si ha

$$A - \text{Id}_3 = \begin{pmatrix} 2 & \frac{1}{2} & 2 \\ -1 & 0 & -1 \\ -1 & -\frac{1}{2} & -1 \end{pmatrix}.$$

Riducendo tale matrice a gradini si trova subito

$$\begin{pmatrix} 2 & \frac{1}{2} & 2 \\ -1 & 0 & -1 \\ -1 & -\frac{1}{2} & -1 \end{pmatrix} \xrightarrow[\substack{R_2 \rightarrow 2R_2 + R_1 \\ R_3 \rightarrow 2R_3 + R_1}]{} \begin{pmatrix} 2 & \frac{1}{2} & 2 \\ 0 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 + R_2} \begin{pmatrix} 2 & \frac{1}{2} & 2 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Quindi il sistema  $(A - \text{Id}_3)x = 0$  si riduce a

$$\begin{cases} 2x_1 + \frac{1}{2}x_2 + 2x_3 = 0 \\ \frac{1}{2}x_2 = 0 \end{cases}$$

La seconda equazione ci dice che  $x_2 = 0$ ; posto  $x_3 = t$ , dalla prima equazione si trova  $2x_1 + 2t = 0$ , ovvero  $x_1 = -t$ . Le soluzioni del sistema, ovvero gli autovettori di  $A$  relativi all'autovalore  $\lambda = 1$ , sono quindi tutte le terne del tipo

$$(-t, 0, t)$$

al variare di  $t \in \mathbb{R}$ . Si noti quindi che benché la molteplicità algebrica di  $\lambda = 1$  sia uguale a 2, la sua molteplicità geometrica è uguale a 1 (in ogni caso, minore, come prevede il Teorema [7.10](#)).

Completiamo l'esempio ricavando i restanti autovettori: per  $\lambda = 2$  si ha invece

$$A - 2\text{Id}_3 = \begin{pmatrix} 1 & \frac{1}{2} & 2 \\ -1 & -1 & -1 \\ -1 & -\frac{1}{2} & -2 \end{pmatrix}.$$

Riduciamo a gradini:

$$\begin{pmatrix} 1 & \frac{1}{2} & 2 \\ -1 & -1 & -1 \\ -1 & -\frac{1}{2} & -2 \end{pmatrix} \xrightarrow[\substack{R_2 \rightarrow R_2 + R_1 \\ R_3 \rightarrow R_3 + R_1}]{} \begin{pmatrix} 1 & \frac{1}{2} & 2 \\ 0 & -\frac{1}{2} & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Quindi il sistema  $(A - 2\text{Id}_3)x = 0$  si riduce a

$$\begin{cases} x_1 + \frac{1}{2}x_2 + 2x_3 = 0 \\ -\frac{1}{2}x_2 + x_3 = 0 \end{cases}.$$

Posto  $x_3 = t$ , dalla seconda equazione si trova  $-\frac{1}{2}x_2 + t = 0$ , ovvero  $x_2 = 2t$ ; sostituendo nella prima, si ha  $x_1 + t + 2t = 0$ , da cui  $x_1 = -3t$ . Le soluzioni del sistema, ovvero gli autovettori di  $A$  relativi all'autovalore  $\lambda = 2$ , sono quindi tutte le terne del tipo

$$(-3t, 2t, t)$$



Dividendo anche la prima riga per 3, si vede che il sistema  $(A - \text{Id}_3)x = 0$  si riduce a

$$\begin{cases} x_1 + x_2 - x_3 = 0 \\ -2x_2 = 0 \end{cases} .$$

La seconda equazione ci dice che  $x_2 = 0$ ; posto  $x_3 = t$ , dalla prima equazione si trova  $x_1 - t = 0$ , ovvero  $x_1 = t$ . Le soluzioni del sistema, ovvero gli autovettori di  $A$  relativi all'autovalore  $\lambda = 1$ , sono quindi tutte le terne del tipo

$$(t, 0, t), \quad t \in \mathbb{R}.$$

Per  $\lambda = 2$  si ha

$$A - 2\text{Id}_3 = \begin{pmatrix} 2 & 3 & -3 \\ -3 & -6 & 3 \\ -1 & -3 & 0 \end{pmatrix} \xrightarrow[\substack{R_2 \rightarrow 2R_2 + 3R_1 \\ R_3 \rightarrow 2R_3 + R_1}]{\longrightarrow} \begin{pmatrix} 2 & 3 & -3 \\ 0 & -3 & -3 \\ 0 & -3 & -3 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - R_2} \begin{pmatrix} 2 & 3 & -3 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

Dividendo anche la seconda riga per -3, si vede che il sistema  $(A - 2\text{Id}_3)x = 0$  si riduce a

$$\begin{cases} 2x_1 + 3x_2 - 3x_3 = 0 \\ x_2 + x_3 = 0 \end{cases}$$

posto  $x_3 = t$ , la seconda equazione dice che  $x_2 = -t$ , e sostituendo nella prima si trova  $2x_1 - 3t - 3t = 0$ , da cui  $x_1 = 3t$ . Le soluzioni del sistema, ovvero gli autovettori di  $A$  relativi all'autovalore  $\lambda = 2$ , sono quindi tutte le terne del tipo

$$(3t, -t, t), \quad t \in \mathbb{R}.$$

Infine, per  $\lambda = -1$  si ha

$$A + \text{Id}_3 = \begin{pmatrix} 5 & 3 & -3 \\ -3 & -3 & 3 \\ -1 & -3 & 3 \end{pmatrix} \xrightarrow[\substack{R_2 \rightarrow 5R_2 + 3R_1 \\ R_3 \rightarrow 5R_3 + R_1}]{\longrightarrow} \begin{pmatrix} 5 & 3 & -3 \\ 0 & -6 & 6 \\ 0 & -12 & 12 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - 2R_2} \begin{pmatrix} 5 & 3 & -3 \\ 0 & -6 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

Dividendo anche la seconda riga per 6, si vede che il sistema  $(A + \text{Id}_3)x = 0$  si riduce a

$$\begin{cases} 5x_1 + 3x_2 - 3x_3 = 0 \\ -x_2 + x_3 = 0 \end{cases}$$

posto  $x_3 = t$ , la seconda equazione dice che  $x_2 = t$ , e sostituendo nella prima si trova  $5x_1 + 3t - 3t = 0$ , da cui  $x_1 = 0$ . Le soluzioni del sistema, ovvero gli autovettori di  $A$

relativi all'autovalore  $\lambda = -1$ , sono quindi tutte le terne del tipo

$$(0, t, t), \quad t \in \mathbb{R}.$$

*Osservazione 7.14.* Data una matrice le cui entrate appartengono a un certo campo, può capitare che le soluzioni dell'equazione caratteristica siano fuori da quel campo. Ad esempio, la matrice a entrate reali

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

ha come equazione caratteristica

$$\det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = 0,$$

che non ha soluzioni nel campo reale  $\mathbb{R}$ , mentre nel campo complesso  $\mathbb{C}$  ha le due soluzioni  $+i$  e  $-i$ .

Quindi, *vista come matrice reale*  $A$  non ha autovalori e quindi neanche autovettori, mentre *vista come matrice complessa* (i numeri reali sono particolari numeri complessi)  $A$  ha i due autovalori  $i$  e  $-i$ , e possiamo determinarne i corrispondenti autovettori, che saranno elementi di  $\mathbb{C}^2$ .

Più in dettaglio, per  $\lambda = i$  si ha

$$A - i\text{Id}_2 = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \xrightarrow{R_2 \rightarrow iR_2 - R_1} \begin{pmatrix} -i & 1 \\ 0 & 0 \end{pmatrix}$$

quindi gli autovettori sono le soluzioni dell'unica equazione  $-ix_1 + x_2 = 0$ . Ponendo  $x_2 = t$  (dove  $t$  stavolta varia tra tutti i numeri complessi) si ha  $x_1 = \frac{1}{i}t$  e quindi gli autovettori relativi all'autovalore  $i$  sono dati da tutte le coppie del tipo  $(\frac{1}{i}t, t)$ , al variare di  $t \in \mathbb{C}$ .

Analogamente, per  $\lambda = -i$  si ha

$$A + i\text{Id}_2 = \begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \xrightarrow{R_2 \rightarrow iR_2 + R_1} \begin{pmatrix} i & 1 \\ 0 & 0 \end{pmatrix}$$

quindi gli autovettori sono le soluzioni dell'unica equazione  $ix_1 + x_2 = 0$ . Ponendo  $x_2 = t$  (dove  $t$  stavolta varia tra tutti i numeri complessi) si ha  $x_1 = -\frac{1}{i}t$  e quindi gli autovettori relativi all'autovalore  $-i$  sono dati da tutte le coppie del tipo  $(-\frac{1}{i}t, t)$ , al variare di  $t \in \mathbb{C}$ .

## 7.3 Matrici diagonalizzabili

La molteplicità geometrica, definita sopra come il numero di parametri che servono per descrivere l'insieme degli autovettori relativi a un dato autovalore, ci dà in effetti il massimo numero di autovettori indipendenti che ogni autovalore fornisce.

Ad esempio, nell'Esempio 7.6 abbiamo visto che gli autovettori di  $A$  relativi all'autovalore  $\lambda = 3$  sono tutti e soli i vettori della forma  $(\frac{-t+s}{2}, t, s)$ . Decomponendo tale vettore come somma della parte che contiene  $t$  e di quella che contiene  $s$

$$\left(\frac{-t+s}{2}, t, s\right) = \left(\frac{-t}{2}, t, 0\right) + \left(\frac{s}{2}, 0, s\right) = t\left(\frac{-1}{2}, 1, 0\right) + s\left(\frac{1}{2}, 0, 1\right)$$

vediamo che tutti gli autovettori sono combinazione dei due vettori indipendenti  $(-\frac{1}{2}, 1, 0)$  e  $(\frac{1}{2}, 0, 1)$ , quindi non troviamo più di due autovettori indipendenti relativi a  $\lambda = 3$ .

Analogamente, sempre nell'Esempio 7.6, gli autovettori di  $A$  relativi all'autovalore  $\lambda = 6$  sono tutte e sole le terne del tipo  $(t, 2t, t)$ . Poiché

$$(t, 2t, t) = t(1, 2, 1)$$

vediamo che tale autovalore fornisce sostanzialmente un solo autovettore assieme a tutti i suoi multipli (quindi non esistono due autovettori indipendenti relativi a tale autovalore).

In effetti, quanto appena visto vale in generale:

**Proposizione 7.15.** *Sia  $A$  una matrice quadrata di ordine  $n$ . Una combinazione lineare*

$$v = \sum_{i=1}^k c_i v_i$$

*di autovettori di  $A$  è un autovalore di  $A$  relativo all'autovalore  $\lambda$  se e solo se tutti gli autovettori  $v_i$  sono relativi allo stesso autovalore  $\lambda$ . In particolare, autovettori relativi ad autovalori distinti sono indipendenti.*

*Dimostrazione.* Dimostriamo prima la sufficienza combinando le due seguenti proprietà.

Siano  $v_1, \dots, v_k$  autovettori di  $A$  relativi all'autovalore  $\lambda$  e sia

$$v = \sum_{i=1}^k c_i v_i$$

per certi coefficienti  $c_i \in \mathbb{K}$ . Allora vogliamo dimostrare che anche  $v$  è un autovettore relativo a  $\lambda$ . In effetti si ha

$$A(v) = A\left(\sum_{i=1}^k c_i v_i\right) = \sum_{i=1}^k c_i A(v_i) = \lambda \sum_{i=1}^k c_i v_i = \lambda v \quad (7.11)$$

dove abbiamo usato le proprietà del prodotto di matrici e del prodotto di uno scalare per una matrice viste nel capitolo precedente. Quindi anche  $v$  è autovettore di  $A$  relativo all'autovalore  $\lambda$ . Per quanto riguarda la necessità, possiamo considerare la somma senza coefficienti e gli autovalori tutti distinti per quanto appena dimostrato. Dobbiamo quindi dimostrare che un autovettore  $v$  relativo a  $\lambda$  è una somma

$$v = \sum_{i=1}^k v_i$$

di autovettori  $v_i$  relativi ad autovalori  $\lambda_i$  solo se  $\lambda = \lambda_i$  per ogni  $i$ . Procediamo per induzione su  $k$ . Quando  $k = 1$  si ha

$$v = v_1 \implies A(v) = A(v_1) \implies \lambda v = \lambda_1 v_1 \implies \lambda = \lambda_1.$$

Ora supponiamo che questo valga per la somma di  $k - 1$  autovettori e dimostriamo che deve valere anche per la somma di  $k$  autovettori. Sia dunque per assurdo

$$v = \sum_{i=1}^k v_i$$

dove  $v, v_1, \dots, v_k$  sono autovettori relativi agli autovalori  $\lambda, \lambda_1, \dots, \lambda_k$ . Allora moltiplicando per  $A$  si ottiene

$$\lambda v = \sum_{i=1}^k \lambda_i v_i$$

e sostituendo la scrittura di  $v$  come somma si ha

$$0 = \sum_{i=1}^k (\lambda_i - \lambda) v_i.$$

Poiché per ipotesi induttiva i  $v_i$  sono linearmente indipendenti, si deve avere  $\lambda_i - \lambda = 0$ , e quindi  $\lambda_i = \lambda$ , per ogni  $i = 1, \dots, k$ .  $\square$

Quindi, nell'Esempio [7.6](#) mettendo insieme gli autovettori indipendenti forniti da ogni autovalore otteniamo in tutto tre autovettori indipendenti  $(\frac{-1}{2}, 1, 0)$ ,  $(\frac{1}{2}, 0, 1)$ ,  $(1, 2, 1)$ , tanti quanti l'ordine della matrice.

**Definizione 7.16.** Le matrici  $A \in M_n(\mathbb{K})$  per cui esistono  $n$  autovettori indipendenti si dicono **matrici diagonalizzabili**.

Quindi, la matrice dell'Esempio 7.6 è diagonalizzabile. Di contro la matrice dell'Esempio 7.12 non lo è, in quanto come autovettori abbiamo trovato  $(-t, 0, t)$  (relativi all'autovalore  $\lambda = 1$ ) e  $(-3t, 2t, t)$  (relativi all'autovalore  $\lambda = 2$ ). Poiché  $(-t, 0, t) = t(-1, 0, 1)$  e  $(-3t, 2t, t) = t(-3, 2, 1)$ , ogni autovalore fornisce un solo autovettore indipendente, quindi in tutto troviamo due autovettori indipendenti, mentre la matrice ha ordine tre.

*Osservazione 7.17.* Si noti che per la Proposizione 7.15 e per la definizione di molteplicità geometrica, una matrice è diagonalizzabile se e solo se la somma delle molteplicità geometriche dei suoi autovalori è pari all'ordine della matrice.

Il motivo per cui una matrice con la proprietà detta si dice diagonalizzabile è il seguente. Supponiamo che  $A$  abbia ordine  $n$  e ammetta  $n$  autovettori indipendenti

$$v_1 = (v_{11}, v_{21}, \dots, v_{n1}), v_2 = (v_{12}, v_{22}, \dots, v_{n2}), \dots, v_n = (v_{1n}, v_{2n}, \dots, v_{nn})$$

Disponendo tali autovettori in colonna, si ottiene una matrice

$$V = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix}$$

quadrata di ordine  $n$  le cui colonne, che sono gli autovettori trovati, sono per ipotesi indipendenti.

Quindi il suo rango (ricordiamo che il rango si può calcolare indifferentemente per righe o per colonne) è  $n$  e per il Teorema 6.11  $V$  è invertibile.

Affermiamo che  $V^{-1}AV$  è una matrice diagonale e più precisamente vale l'uguaglianza

$$\begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix}^{-1} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad (7.12)$$

cioè

$$V^{-1}AV = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

dove  $\lambda_1$  è l'autovalore di  $v_1$ ,  $\lambda_2$  è l'autovalore di  $v_2$ , e così via.

Infatti, moltiplicando entrambi i membri di tale uguaglianza a sinistra per  $V$ , vediamo

che essa equivale a

$$AV = V \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad (7.13)$$

Per mostrare l'uguaglianza scritta in questa forma, osserviamo che, per definizione di prodotto righe per colonne, la prima colonna del prodotto di matrici al primo membro della (7.13) si ottiene moltiplicando le righe di  $A$  per la prima colonna di  $V$ , ovvero

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix}$$

Ma, essendo la prima colonna di  $V$  l'autovettore  $v_1$  di  $A$ , si ha

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix} = \lambda_1 \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix} = \begin{pmatrix} \lambda_1 v_{11} \\ \lambda_1 v_{21} \\ \vdots \\ \lambda_1 v_{n1} \end{pmatrix}$$

Facendo lo stesso ragionamento sulle altre colonne del prodotto di matrici al primo membro della (7.13), si ottiene allora

$$AV = \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} & \cdots & \lambda_n v_{1n} \\ \lambda_1 v_{21} & \lambda_2 v_{22} & \cdots & \lambda_n v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 v_{n1} & \lambda_2 v_{n2} & \cdots & \lambda_n v_{nm} \end{pmatrix} \quad (7.14)$$

Questa, come si verifica subito sempre svolgendo un prodotto righe per colonne, è uguale al secondo membro della (7.13).

*Esempio 7.18.* Consideriamo la matrice

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

di cui abbiamo calcolato autovalori e autovettori nell'Esempio 7.5 trovando  $(t, t)$  come autovettori relativi all'autovalore  $\lambda = 4$  e  $(-t, t)$  come autovettori relativi all'autovalore  $\lambda = -2$ . Poiché  $(t, t) = t(1, 1)$  e  $(-t, t) = t(-1, 1)$ , troviamo due autovettori indipendenti  $(1, 1)$  e  $(-1, 1)$  e la matrice è diagonalizzabile. Per verificare la validità della

(7.12), disponiamo tali autovettori in colonna, ottenendo

$$V = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Come si verifica subito, l'inversa di tale matrice è

$$V^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

e si ha

$$V^{-1}AV = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 4 & 2 \\ 4 & -2 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$$

cioè, come previsto dalla (7.12), otteniamo la matrice diagonale che ha sulla diagonale gli autovalori di  $A$ .

La diagonalizzazione è fondamentale quando si vuole cambiare sistema di riferimento in modo da passare a un nuovo sistema in cui il problema in esame assume una forma più semplice. Ad esempio, se  $A$  rappresenta una superficie quadrica come nel secondo esempio visto nel primo paragrafo, la diagonalizzazione permette di ottenere un nuovo sistema di riferimento in cui l'equazione della superficie è la più semplice possibile.

## 7.4 Applicazioni alla grafica: rotazioni nello spazio e stitching di immagini

In quest'ultimo paragrafo, accenneremo ad alcune applicazioni di natura grafica della teoria di autovalori e autovettori.

La prima applicazione che vedremo consiste nella determinazione dell'asse e dell'angolo di una rotazione data nello spazio tridimensionale. Prima di entrare nei dettagli, vediamo alcune generalità su tali trasformazioni.

Come abbiamo visto nella Sezione 6.1, fissato nello spazio un sistema di riferimento con origine  $O$ , le applicazioni lineari sono descritte mediante una matrice  $A$  di ordine 3. Più precisamente se  $x$  rappresenta la terna delle coordinate di un punto  $P$  (o del corrispondente vettore  $\vec{OP}$ ), le coordinate del punto (o del vettore) trasformato sono date dal prodotto  $Ax$ .

Tra le applicazioni lineari che abbiamo visto, rotazioni e riflessioni hanno un'altra caratteristica in comune (che ad esempio le proiezioni non hanno). Queste sono infatti trasformazioni rigide, ovvero non modificano né le lunghezze dei vettori né gli angoli tra i vettori.

Ricordiamo che nel Capitolo 4 abbiamo visto come, se il sistema di riferimento fissato è ortonormale, allora la lunghezza di un vettore  $\vec{OP}$  di coordinate  $x = (x_1, x_2, x_3)$  è dato da  $|x| = \sqrt{x_1^2 + x_2^2 + x_3^2}$ , cf. (4.20). Inoltre l'angolo  $\theta$  tra due vettori  $\vec{OP}$  e  $\vec{OQ}$  di coordinate rispettivamente  $x = (x_1, x_2, x_3)$  e  $y = (y_1, y_2, y_3)$  è dato dalla formula

$$\cos(\theta) = \frac{x \cdot y}{|x||y|},$$

dove a numeratore abbiamo il prodotto scalare  $x \cdot y = x_1y_1 + x_2y_2 + x_3y_3$ , cf. (4.21). Osservando che anche la lunghezza  $|x|$  di un vettore si scrive in termini del prodotto scalare, in quanto  $\sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{x \cdot x}$ , deduciamo che una trasformazione è rigida, ovvero non modifica né lunghezze né angoli, se e solo se non modifica il prodotto scalare.

In formula, dal momento che le coordinate del vettore  $x$  dopo la trasformazione sono date da  $Ax$ , la trasformazione rappresentata dalla matrice  $A$  sarà una trasformazione rigida se e solo se

$$(Ax) \cdot (Ay) = x \cdot y \quad (7.15)$$

qualunque siano le terne  $x$  e  $y$ .

Ora, come abbiamo visto a pagina 241 il prodotto scalare  $x \cdot y$  può essere anche scritto come prodotto  $x^T y$  tra la trasposta della colonna formata dalle componenti di  $x$  e la colonna delle componenti di  $y$ . Questo ci permette di riscrivere la condizione (7.15) come

$$(Ax)^T (Ay) = x^T y. \quad (7.16)$$

Il vantaggio di questa riformulazione è che possiamo applicare la proprietà (2) della trasposta, per cui  $(Ax)^T = x^T A^T$ . Quindi la (7.16) si riscrive

$$x^T A^T A y = x^T y. \quad (7.17)$$

Ora, questa uguaglianza può essere verificata per ogni  $x$  e ogni  $y$  solo se il prodotto  $A^T A$  che compare al primo membro è uguale alla matrice identica  $\text{Id}_3$ .

In altre parole, una trasformazione lineare è rigida se e solo se la matrice che la rappresenta rispetto a un sistema di riferimento ortonormale fissato è tale che  $A^T A = \text{Id}_3$ .

**Definizione 7.19.** Una matrice quadrata di ordine  $n$  a entrate reali  $A \in M_n(\mathbb{R})$  si dice **ortogonale** se  $A^T A = \text{Id}_n$  (equivalentemente, se  $A^{-1} = A^T$ ).

*Osservazione 7.20.* In generale, il determinante di una matrice ortogonale può essere

uguale solo a  $+1$  o  $-1$ . Infatti, dalla condizione  $A^T A = \text{Id}_n$  abbiamo che

$$\det(A^T A) = \det(\text{Id}_n) = 1$$

ovvero, applicando il teorema di Binet al primo membro

$$\det(A^T) \det(A) = 1.$$

Poiché, come abbiamo visto nella proprietà (4), il determinante di una matrice è uguale a quello della sua trasposta, l'ultima uguaglianza equivale a

$$\det(A)^2 = 1$$

da cui deduciamo che  $\det(A) = \pm 1$ , come affermato.

Alla luce di questa definizione, possiamo riassumere quanto detto finora affermando che una trasformazione lineare è rigida se e solo se la matrice che la rappresenta rispetto a un sistema di riferimento ortonormale fissato è una matrice ortogonale.

Quindi la matrice che rappresenta una rotazione è sicuramente ortogonale, ma ovviamente non vale il viceversa. Ad esempio, anche le riflessioni sono rappresentate da matrici ortogonali in quanto sono trasformazioni rigide. Tuttavia, per riconoscere le rotazioni tra le matrici ortogonali basta calcolare un determinante, come afferma il seguente

**Teorema 7.21.** *Fissato un sistema di riferimento ortonormale nello spazio, una matrice  $A$  quadrata di ordine 3 rappresenta una rotazione se e solo se  $A$  è ortogonale e  $\det(A) = +1$ .*

**Definizione 7.22.** L'insieme delle matrici ortogonali di ordine  $n$  si denota con  $O(n)$ . Il sottoinsieme di  $O(n)$  costituito dalle matrici ortogonali di ordine  $n$  e determinante  $+1$  (dette anche **matrici ortogonali speciali**) si denota con  $SO(n)$ .

**Proposizione 7.23.** *Gli insiemi  $O(n)$  e  $SO(n)$  sono sottogruppi di  $GL_n(\mathbb{R})$  detti rispettivamente **gruppo ortonormale** e **gruppo ortonormale speciale** di ordine  $n$ .*

*Dimostrazione.* Incominciamo col dimostrare che  $O(n)$  è un sottogruppo di  $GL_n(\mathbb{R})$  rispetto al prodotto di matrici. Questo discende dalle seguenti proprietà:

- (i) Chiaramente  $\text{Id} \in O(n)$  poiché  $\text{Id}^T = \text{Id}$  e  $\text{Id Id} = \text{Id}$ .
- (ii) Se  $A, B \in O(n)$  allora anche  $AB \in O(n)$ . Questo deriva da

$$(AB)^T (AB) = B^T A^T AB = B^T \text{Id} B = B^T B = \text{Id}$$

dove abbiamo utilizzato la proprietà (2) ed il fatto che  $A, B \in O(n)$ .

(iii) Se  $A \in O(n)$  allora  $A^{-1} \in O(n)$ . Qua usiamo il fatto che  $(A^T)^T = A$  per cui

$$(A^T)^T A^T = AA^T = AA^{-1} = \text{Id}$$

dove la penultima uguaglianza deriva da  $A^T = A^{-1}$ .

Ora dobbiamo mostrare le stesse proprietà per  $SO(n)$ , ossia aggiungere la condizione  $\det(A) = \det(B) = 1$  e richiedere  $\det(\text{Id}) = \det(AB) = \det(A^{-1}) = 1$ .

(i) Chiaramente  $\text{Id} \in SO(n)$  poiché  $\det(\text{Id}) = 1$  e  $\text{Id} \in O(n)$ .

(ii) Se  $A, B \in SO(n)$  allora anche  $AB \in SO(n)$ . Questo deriva da  $AB \in O(n)$  e da

$$\det(AB) = \det(A) \det(B) = 1$$

dove abbiamo utilizzato che il teorema di Binet ed il fatto che  $A, B \in SO(n)$ .

(iii) Se  $A \in SO(n)$  allora  $A^{-1} \in SO(n)$ . Qua usiamo il fatto che  $(A^T) = A^{-1}$  per cui

$$\det(A^{-1}) = \det(A^T) = \det(A) = 1.$$

□

Ora vediamo alcuni esempi concreti di matrici di rotazione.

Iniziamo dal caso del piano, che è il più semplice. In tal caso, fissato un sistema di riferimento ortonormale con origine del sistema data dal punto  $O$ , la matrice che rappresenta la rotazione in senso antiorario di angolo  $\theta$  attorno a  $O$  è determinata da come agisce sui vettori della base  $O\vec{P}_1$  e  $O\vec{P}_2$ . In altre parole le entrate della matrice

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

che rappresenta la rotazione sono determinate dal prodotto per i vettori colonna  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  e  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ . Si può facilmente verificare graficamente che

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$$

da cui ricaviamo  $a_{11} = \cos \theta$  e  $a_{21} = \sin \theta$ . Allo stesso modo da

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$$

si ottiene  $a_{22} = \cos \theta$  e  $a_{12} = -\sin \theta$ .

Per cui un elemento generico di  $SO(2)$ , ovvero una rotazione di un angolo  $\theta$  in senso antiorario attorno all'origine, è dato da

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (7.18)$$

Ovvero la rotazione è rappresentata in coordinate dalla funzione

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos \theta x_1 - \sin \theta x_2 \\ \sin \theta x_1 + \cos \theta x_2 \end{pmatrix} \quad (7.19)$$

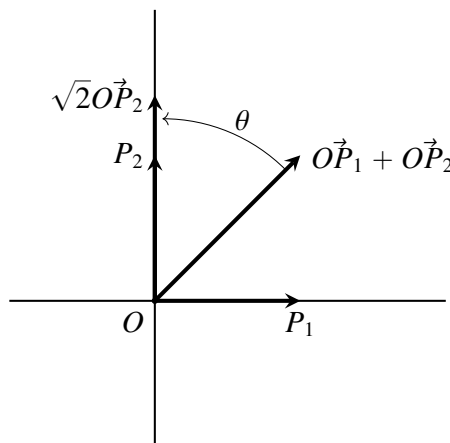
*Esempio 7.24.* Posto  $\theta = \frac{\pi}{4}$ , si ha  $\cos \theta = \sin \theta = \frac{\sqrt{2}}{2}$ . Quindi la rotazione di angolo  $\frac{\pi}{4}$  è rappresentata dalla matrice

$$\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

Ad esempio, il vettore  $\vec{OP}$  del piano di coordinate  $(1, 1)$  rispetto al sistema di riferimento fissato viene mandato dalla rotazione nel vettore di coordinate

$$\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix}$$

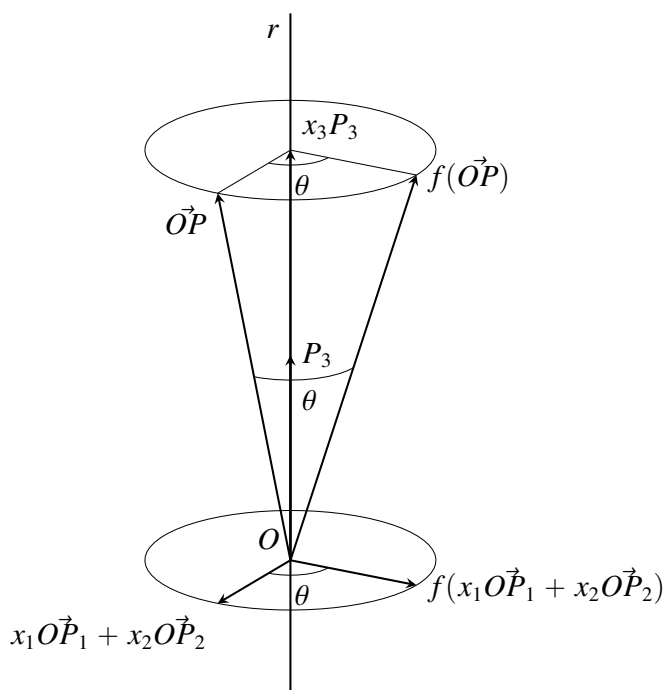
Questo può essere verificato graficamente: se  $\vec{OP}_1, \vec{OP}_2$  sono i due vettori che formano la base del sistema di riferimento fissato, il fatto che  $\vec{OP}$  abbia coordinate  $(1, 1)$  significa che  $\vec{OP} = \vec{OP}_1 + \vec{OP}_2$ , mentre il fatto che il suo trasformato abbia coordinate  $(0, \sqrt{2})$  significa che esso è  $0\vec{OP}_1 + \sqrt{2}\vec{OP}_2 = \sqrt{2}\vec{OP}_2$ . Infatti, come si vede nel seguente disegno,  $\sqrt{2}\vec{OP}_2$  è esattamente il vettore che si ottiene ruotando di 45 gradi in senso antiorario il vettore  $\vec{OP} = \vec{OP}_1 + \vec{OP}_2$ :



Nello spazio tridimensionale, la situazione è un po' più complessa in quanto le rotazioni dello spazio sono determinate non solo dall'angolo di rotazione ma anche dal loro asse, ovvero dalla retta attorno alla quale effettuiamo la rotazione. Per questo motivo, non esiste una classificazione esplicita completa come quella data dalla (7.19) nel caso del piano.

Assumono però forma particolarmente semplice le rotazioni attorno agli *assi coordinati*, ovvero alle rette che contengono i vettori  $\vec{OP}_1, \vec{OP}_2, \vec{OP}_3$  che formano la base del sistema di riferimento scelto: le chiameremo rispettivamente *asse x*, *asse y* e *asse z*.

Iniziamo dalle rotazioni attorno all'asse  $z$ . Come si vede nel seguente disegno, una rotazione di angolo  $\theta$  di un vettore  $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2 + x_3\vec{OP}_3$  attorno a tale asse mantiene fissa la componente  $x_3\vec{OP}_3$  che sta sull'asse di rotazione, mentre ruota di un angolo  $\theta$  la componente  $x_1\vec{OP}_1 + x_2\vec{OP}_2$  che sta nel piano determinato da  $\vec{OP}_1$  e  $\vec{OP}_2$ .



Quindi, applicando la matrice che rappresenta la rotazione alla terna  $(x_1, x_2, x_3)$  delle coordinate del vettore  $\vec{OP}$ , la terza componente  $x_3$  non deve variare mentre le prime due componenti  $x_1$  e  $x_2$  si trasformeranno esattamente come per le rotazioni nel piano, ovvero secondo la (7.19). Deduciamo che la matrice che rappresenta una rotazione di angolo  $\theta$  attorno all'asse  $z$  è

$$\begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7.20)$$

Infatti, come si vede applicando tale matrice a una terna  $(x_1, x_2, x_3)$ ,

$$\begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \cos \theta x_1 - \sin \theta x_2 \\ \sin \theta x_1 + \cos \theta x_2 \\ x_3 \end{pmatrix}$$

come previsto.

Ragionamenti analoghi permettono di dedurre che la matrice di una rotazione di angolo  $\theta$  attorno all'asse  $y$  è

$$\begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \quad (7.21)$$

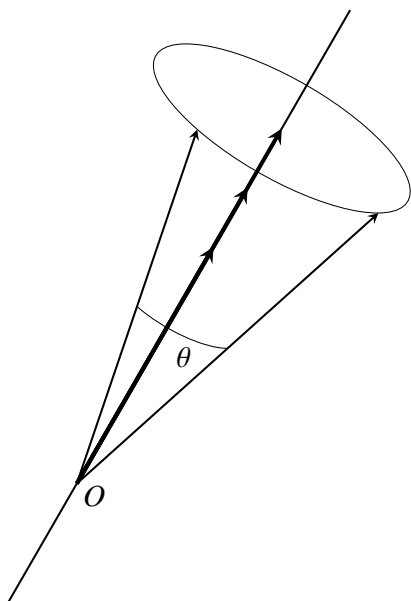
Infatti applicando tale matrice a una terna di coordinate  $(x_1, x_2, x_3)$  non si modifica  $x_2$  che rappresenta la componente corrispondente all'asse di rotazione, mentre  $x_1$  e  $x_3$  si trasformano come una rotazione nel piano determinato da  $\vec{OP}_1$  e  $\vec{OP}_3$ . Analogamente la matrice di una rotazione di angolo  $\theta$  attorno all'asse  $x$  è

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \quad (7.22)$$

Applicando tale matrice a una terna di coordinate  $(x_1, x_2, x_3)$  non si modifica  $x_1$  che rappresenta la componente corrispondente all'asse  $x$  di rotazione, mentre  $x_2$  e  $x_3$  si trasformano come una rotazione nel piano su cui giacciono  $\vec{OP}_2$  e  $\vec{OP}_3$ .

Data una qualunque matrice  $A \in \text{SO}(3)$  possiamo trovare l'asse di rotazione e l'angolo di rotazione come segue.

Per quello che riguarda l'asse, osserviamo che i vettori che appartengono all'asse di una rotazione rimangono invariati quando applichiamo la rotazione.



Poiché applicando la rotazione  $A$  il vettore di coordinate  $x$  viene mandato nel vettore di coordinate  $Ax$ , le coordinate  $x$  dei vettori che stanno sull'asse devono soddisfare l'uguaglianza  $Ax = x$ . Ma questa è esattamente l'uguaglianza che soddisfano gli autovettori di  $A$  relativi all'autovalore 1, e quindi per determinare l'asse basta determinare tali autovettori.

Quanto all'angolo di rotazione, utilizzando l'identificazione tra  $R^2$  e  $\mathbb{C}$  ed il fatto che un'equazione caratteristica ammette sempre 3 soluzioni (in  $\mathbb{C}$ ), si può dimostrare che esso è determinato dalle restanti soluzioni dell'equazione caratteristica. Più precisamente, si mostra che gli autovalori complessi di una matrice in  $SO(2)$  sono numeri complessi del tipo

$$\cos \theta \pm i \sin \theta \quad (7.23)$$

dove  $\theta$  è proprio l'angolo di rotazione. Lo stesso vale per le ulteriori radici complesse del polinomio caratteristico di  $A$ .

Ora, a partire dalle rotazioni date attorno agli assi, possiamo ricavare altre rotazioni per composizione: infatti, il Teorema [7.21](#) e la Proposizione [7.23](#) ci permettono di dedurre che la composizione di due rotazioni nello spazio è ancora una rotazione. Più precisamente, se abbiamo due rotazioni rappresentate da due matrici (ortogonali speciali)  $A_1, A_2 \in SO(n)$ , la loro composizione come sappiamo è rappresentata dal prodotto  $A_1 A_2$ . Per la Proposizione [7.23](#) questa è ancora un elemento di  $SO(n)$  e per il Teorema [7.21](#) questa è una rotazione.

Illustriamo ora con un esempio come determinare, usando il calcolo di autovalori e autovettori, asse e angolo di una rotazione.



Come si vede confrontando con la (7.23), l'angolo di rotazione  $\theta$  soddisfa  $\cos \theta = -\frac{1}{2}$ ,  $\sin \theta = \frac{\sqrt{3}}{2}$ , e concludiamo  $\theta = \frac{2}{3}\pi$

Per trovare l'asse di rotazione, calcoliamo gli autovettori relativi all'autovalore  $\lambda = +1$  risolvendo il sistema omogeneo  $(A - \text{Id}_3)x = 0$ . Si ha

$$A - \text{Id}_3 = \begin{pmatrix} -1 & 0 & -1 \\ -1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - R_1} \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 + R_2} \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Dunque gli autovettori relativi a  $\lambda = 1$  sono le soluzioni del sistema ridotto

$$\begin{cases} -x_1 - x_3 = 0 \\ -x_2 + x_3 = 0 \end{cases}$$

Posto  $x_3 = t$ , si trova  $x_2 = x_3 = t$  e  $x_1 = -x_3 = -t$ ; quindi le soluzioni sono le triple  $(-t, t, t) = t(-1, 1, 1)$ , al variare di  $t \in \mathbb{R}$ . Queste rappresentano in coordinate i vettori  $-t\vec{OP}_1 + t\vec{OP}_2 + t\vec{OP}_3 = t(-\vec{OP}_1 + \vec{OP}_2 + \vec{OP}_3)$  che formano l'asse di rotazione.

Concludiamo la discussione sulle matrici ortogonali enunciando una importante caratterizzazione di queste.

**Proposizione 7.26.** Una matrice  $A \in M_n(\mathbb{R})$  è ortogonale se e solo se le sue colonne (o equivalentemente le sue righe) rappresentano  $n$  vettori ortogonali a due a due.

Chiudiamo il capitolo citando un'importante applicazione della teoria degli autovettori e autovettori all'informatica e in particolare alla cosiddetta computer vision.

Una categoria importante di matrici che soddisfano svariate proprietà è data dalle cosiddette *matrici simmetriche*

**Definizione 7.27.** Una matrice  $A \in M_n(\mathbb{R})$  si dice **simmetrica** se  $A^T = A$ , ovvero se le sue entrate soddisfano  $A_{ij} = a_{ji}$  per ogni coppia di indici.

Ad esempio, la matrice  $A = \begin{pmatrix} 3 & 2 \\ 2 & 5 \end{pmatrix}$  è, come si verifica immediatamente, simmetrica. enunciando ora alcune proprietà delle matrici simmetriche senza dimostrarle tutte:

- (i) Le matrici simmetriche di ordine  $n$  non formano un sottogruppo di  $M_n(\mathbb{K})$ . Infatti il prodotto di due matrici simmetriche non è necessariamente una matrice simmetrica. In effetti possiamo vedere che per due matrici simmetriche  $A$  e  $B$  vale la seguente condizione

$$AB = BA \iff AB \text{ è simmetrica.}$$

Infatti le matrici al primo membro hanno entrate

$$(AB)_{ij} = \sum_{k=1}^n a_{ik}b_{kj} \text{ e } (BA)_{ij} = \sum_{k=1}^n b_{ik}a_{kj}.$$

Scambiando gli indici delle entrate di  $A$  e  $B$  vediamo queste due espressioni si equivalgono se e solo se  $AB$  è una matrice simmetrica.

- (ii) Una matrice simmetrica ammette sempre  $n$  autovalori reali (possibilmente con ripetizioni) ed  $n$  autovettori linearmente indipendenti.
- (iii) Gli autovettori di una matrice simmetrica  $A$  relativi ad autovalori distinti sono ortogonali. Infatti siano  $v_1$  e  $v_2$  autovettori relativi rispettivamente a  $\lambda_1$  e  $\lambda_2$  con  $\lambda_1 \neq \lambda_2$ . Allora vogliamo mostrare che il prodotto scalare  $v_1 \cdot v_2$  è nullo. Questo discende dal calcolo

$$\lambda_1 v_1 \cdot v_2 = \lambda_1 v_1^T v_2 = (Av_1)^T v_2 = v_1^T A^T v_2 = v_1^T A v_2 = \lambda_2 v_1^T v_2 = \lambda_2 v_1 \cdot v_2$$

dove abbiamo usato che  $A$  è una matrice simmetrica. Ora l'uguaglianza può valere solo se  $v_1 \cdot v_2 = 0$  poiché  $\lambda_1 \neq \lambda_2$  per assunto.

Abbiamo detto sopra che se una matrice  $A$  è diagonalizzabile (cioè ammette  $n$  autovettori indipendenti), esiste sempre una matrice invertibile  $V$ , le cui colonne sono gli autovettori di  $A$ , tale che  $V^{-1}AV$  sia una matrice diagonale (avente sulla diagonale gli autovalori di  $A$ ). Per quanto appena visto, se  $A$  è simmetrica questo è sempre possibile e, per la Proposizione [7.26](#),  $V$  è una matrice ortogonale. A meno di scambiare due colonne possiamo supporre che  $V \in SO(n)$  per cui abbiamo il seguente

**Teorema 7.28.** *Sia  $A \in M_n(\mathbb{R})$  una matrice simmetrica con autovalori  $\lambda_1, \dots, \lambda_n$ . Allora esiste una matrice ortogonale speciale  $V \in SO(n)$  tale che*

$$V^{-1}AV = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Denotando con  $D$  la matrice diagonale  $V^{-1}AV$ , l'uguaglianza  $V^{-1}AV = D$  si può anche scrivere, moltiplicando entrambi i membri a destra per  $V^{-1}$  e a sinistra per  $V$  come  $A = VDV^{-1}$ .

Se siamo nel caso di matrici di ordine 3, tale uguaglianza ci dice che ogni trasformazione dello spazio che sia rappresentata da una matrice simmetrica può essere decomposta come prodotto (quindi composizione delle trasformazioni corrispondenti) di una rotazione, poi di una matrice diagonale e poi della rotazione inversa alla prima.

Tale risultato trova applicazione ad esempio in alcuni algoritmi di “stitching di immagini”, ovvero il procedimento mediante il quale si combinano più fotografie, sovrappo-  
nendole, per ottenere un’unica fotografia panoramica o in un video stabilizzato in caso  
si tratti di frame di un video. Il riconoscimento delle parti corrispondenti nelle varie  
immagini che devono essere incollate usa il risultato di decomposizione delle matrici  
simmetriche che abbiamo appena citato, ma non entriamo nei dettagli.